

AD\_\_\_\_\_

Award Number DAMD17-98-1-8086

TITLE: Cloning of an ETS-Related Transcription Factor Involved in  
a Novel Epigenetic Mechanism of Mammary Carcinogenesis

PRINCIPAL INVESTIGATOR: Helmut Zarbl, Ph.D.

CONTRACTING ORGANIZATION: Fred Hutchinson Cancer Research Center  
Seattle, Washington 98104-1024

REPORT DATE: May 1999

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command  
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;  
Distribution Unlimited

The views, opinions and/or findings contained in this report are  
those of the author(s) and should not be construed as an official  
Department of the Army position, policy or decision unless so  
designated by other documentation.

20000829 039

| REPORT DOCUMENTATION PAGE   |   |  | Form Approved<br>OMB No. 0704-0188                  |  |
|---|---|--|---|--|
| Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.  |   |  |   |  |
| 1. AGENCY USE ONLY (Leave blank)  |   | 2. REPORT DATE<br>May 1999                                 |   | 3. REPORT TYPE AND DATES COVERED<br>Annual (15 Apr 98 - 14 Apr 99) |
| 4. TITLE AND SUBTITLE<br>Cloning of an ETS-Related Transcription Factor Involved in a Novel Epigenetic Mechanism of Mammary Carcinogenesis  |   |  | 5. FUNDING NUMBERS<br>DAMD17-98-1-8086              |  |
| 6. AUTHOR(S)<br>Helmut Zarbl, Ph.D.   |   |  |   |  |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)<br>Fred Hutchinson Cancer Research Center<br>Seattle, Washington 98104-1024  |   |  | 8. PERFORMING ORGANIZATION<br>REPORT NUMBER         |  |
| 9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)<br>U.S. Army Medical Research and Materiel Command<br>Fort Detrick, Maryland 21702-5012   |   |  | 10. SPONSORING / MONITORING<br>AGENCY REPORT NUMBER |  |
| 11. SUPPLEMENTARY NOTES   |   |  |   |  |
| 12a. DISTRIBUTION / AVAILABILITY STATEMENT<br>Approved for Public Release; Distribution Unlimited   |   |  | 12b. DISTRIBUTION CODE                              |  |
| 13. ABSTRACT (Maximum 200 words)      Previous studies demonstrated that activating <i>Hras1</i> mutations found in NMU-induced rat mammary tumors arise as background mutations within the developing gland, and that carcinogenic doses of NMU induce the tumorigenic outgrowth of these mutants via epigenetic mechanisms. Analysis of DNA sequence comprising the <i>Hras1</i> promoter revealed a mammary-specific DNA conformation (CS) which was disrupted by NMU exposure. When we compared the binding of proteins among cell types to an <i>ets</i> -like response element present within the CS region, we detected the formation of mammary specific DNA complexes. Using sequence specific DNA affinity chromatography, we purified micrograms of two proteins with estimated molecular weights of ~42-43 and ~37-38 kDa that are present in mammary cells and bind specifically to this promoter element. Amino acid sequencing by tandem mass spectrometry indicated that both proteins include peptides corresponding to highly conserved sequences present in a family of transcription factors. These finding suggest that carcinogen disruption of the CS could permit binding of transcription factors such as CBF-A and lead to deregulation <i>Hras1</i> expression. Characterization of these binding proteins and their role in regulation of tissue specific gene expression should provide insight into epigenetic mammary carcinogenesis in rats and humans. |   |  |   |  |
| 14. SUBJECT TERMS<br>Breast Cancer<br><br>transcription factors, epigenetic carcinogenesis, DNA conformation,, <i>Hras</i> ,  |   |  | 15. NUMBER OF PAGES<br><br>21                       |  |
|   |   |  | 16. PRICE CODE                                      |  |
| 17. SECURITY CLASSIFICATION<br>OF REPORT<br>Unclassified  | 18. SECURITY CLASSIFICATION<br>OF THIS PAGE<br>Unclassified | 19. SECURITY CLASSIFICATION<br>OF ABSTRACT<br>Unclassified | 20. LIMITATION OF ABSTRACT<br>Unlimited             |  |

## FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

\_\_\_\_ Where copyrighted material is quoted, permission has been obtained to use such material.

\_\_\_\_ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

\_\_\_\_ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

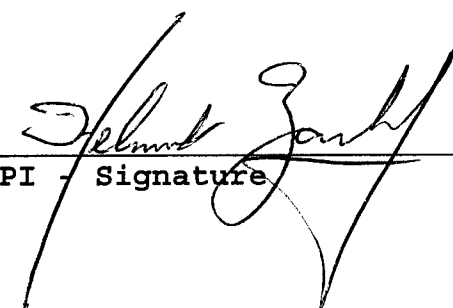
✓ \_\_\_\_ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

\_\_\_\_ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

\_\_\_\_ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

\_\_\_\_ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

\_\_\_\_ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

  
PI - Signature \_\_\_\_\_ Date MAY 11/99

## Table of Contents

|   |    |
|---|----|
| Front Cover .....                                 | 1  |
| Report Documentation Page .....                   | 2  |
| Foreword .....                                    | 3  |
| Table of Contents .....                           | 4  |
| Introduction .....                                | 5  |
| Body of Report .....                              | 6  |
| Methods.....                                      | 6  |
| Cell culture and protein purification.....        | 6  |
| Electrophoretic Mobility Shift Assays (EMSA)..... | 7  |
| Amino acid analysis.....                          | 8  |
| Results.....                                      | 9  |
| Discussion .....                                  | 10 |
| Future considerations .....                       | 11 |
| Key Research Accomplishments .....                | 11 |
| Reportable Outcomes .....                         | 12 |
| Conclusions .....                                 | 12 |
| References .....                                  | 13 |
| Appendices .....                                  | 14 |
| Table 1 .....                                     | 14 |
| Figure 1 .....                                    | 15 |
| Figure 2 .....                                    | 16 |
| Figure 3 .....                                    | 17 |
| Figure 4 .....                                    | 18 |
| Figure 5 .....                                    | 19 |
| Figure 6 .....                                    | 20 |
| Figure 7 .....                                    | 21 |

## Introduction

Our previous finding suggested a new paradigm for non-mutagenic mammary carcinogenesis (1). We demonstrated that the activating *Hras1* mutations found in NMU-induced tumors arise as background mutations within cells of the developing gland, and that NMU enhances the phenotypic penetrance of these mutations by initiating alterations in DNA conformation. We subsequently identified a cell-type-specific, DNA structure or conformational switch (CS) within the *Hras1* promoter of normal rat mammary cells (RMCs) *in vivo* (2). The analogous structure was also detected in the promoter of the human *Hras* gene. Our results further demonstrated that depending upon hormonal status of the animals, RMCs switch between states where the structural feature is present or absent from the *Hras1* promoter. In the F344 and SD rat strains, which are sensitive to mammary carcinogenesis, a carcinogenic dose of NMU initiated the loss of this structure from the *Hras1* promoter of RMCs. By contrast, the same exposure failed to disruption the promoter structure in Cop rats, which are highly resistant to mammary carcinogenesis. NMU also failed to promote the outgrowth of pre-existing *Hras1* mutants present in the mammary epithelium of resistant Cop rats. Phenotypic analysis of resistant (F344 X Cop)F1 progeny further indicated that the suppression of CS disruption was mediated by one or more suppressors expressed in RMCs of Cop rats. Together our results suggest that NMU-induced alterations in DNA conformation promote the outgrowth of pre-existing mutants by irreversibly deregulating expression of *Hras1*, thereby increasing the phenotypic penetrance of conditional oncomutations. Analysis of the DNA sequence comprising the CS region suggested the presence of an *ets*-like transcription factor responsive element. When we compared the binding of proteins from a variety of cell types to this response element, we detected the formation of a major complex with mammary cell extracts that was distinct from that formed in other cell types. Together these results supported a model in which CS disruption unmask the *ets*-like responsive element contained within the CS structure, and thereby allows increased binding of transcription factors and elevated expression of responsive genes. The overall goals of these studies are to i) identify and clone the mammary cells specific transcription factor that forms a complex with the *ets*-like responsive element present in the CS switch region, and ii) determine its role in mammary carcinogenesis.

## Body of Report

The technical objective approved for the first 12 months of this study was the purification and amino acid sequencing of the transcription factor that binds specifically to the *ets*-like response element present in the CS region of the *Hras1* promoter. These objectives were successfully completed, and the results obtained are detailed below.

We previously demonstrated that the complex formed with nuclear extracts from mammary epithelial cells was distinct from those formed with nuclear extracts from other cell types. These results indicated that either the proteins cell proteins capable of interacting with the CS region were distinct from those in other cell types or undergo mammary cell specific post-translational modifications. In either case it was essential that the protein(s) binding to the CS region be cloned from mammary epithelial cells. In the original proposal we outlined several alternative strategies for cloning the gene that encodes the transcription factor that binds to the *Hras1* promoter CS region. However, given the high affinity of the specific binding protein for the double-stranded oligonucleotide encoding the CS region, purification of the proteins by affinity chromatography followed by amino acid sequencing was the method of choice. Our previous results had indicated that presence of the mammary cell specific activity in nuclear extracts from the BICR cell line, which was derived from a mammary tumor induced in a Fischer 344 rat by exposure to the carcinogen dimethylbenz(a)anthracene (DMBA). The BICR cell line could easily be grown in large quantities was therefore used as the source of the nuclear extract for purification of the transcription factor.

## Methods

### Cell culture and protein purification.

BICR rat mammary tumor cells were grown on 15 centimeter dishes using DMEM culture medium supplemented with 5% calf serum (HyClone Laboratories, Inc, UT). Since our previous studies indicated that the CS binding activity was cell cycle-dependent, we harvested cells during the exponential growth (80-90 % confluence). Cells were harvested by scraping into PBS and collected by centrifugation at 800 x g. Cells were lysed in buffer A (10 mM HEPES pH7.9, 10 mM KCl, 1.5 mM MgCl<sub>2</sub>, 0.2% Triton X-100, 0.5 mM DTT) for 30 min. and passed through 26G needle to break cell membranes (3). Intact nuclei were collected by centrifugation at 500 x g, resuspended and incubated in a buffer C (10 mM HEPES pH7.9, 0.42 M NaCl, 1.5 mM MgCl<sub>2</sub>, 0.2 mM EDTA, 0.5 mM DTT, 0.5 mM PMSF, 20% Glycerol) for 45 min with rotation. Nuclear proteins were dialyzed against binding buffer D (20 mM HEPES, pH7.9, 10% glycerol, 0.1 M KCl, 0.2 M EDTA, 0.5 mM DTT, 0.5 mM PMSF) and centrifuged at 30,000 x g to

remove insoluble components. Nuclear proteins were extracted from several different batches of cell preparations. Each batch of extracted proteins was tested for presence of binding activity using the electrophoretic mobility shift assay (EMSA) initially used to detect the binding activity prior to affinity purification (see protocol below). The combined wet volume of the cell pellets used for protein purification was approximately ~35 ml.

Affinity chromatography was performed essentially as described by Kadonaga (4). Biotinylated sense and antisense oligonucleotides corresponding the wild-type *ets*-like response element and the M2 mutant sequence that fails to bind the mammary specific transcription factor (Figure 1) were obtained from Research Genetics. The corresponding pairs of oligomers were annealed and attached to streptavidin-coated agarose beads (Pierce). The individual batches of nuclear protein extract were first incubated with poly dI•dC at 50 ug/ml to titrate non-specific DNA binding proteins and centrifuged at 12,000 x g to remove precipitates. To further reduce to amount of non-specific DNA binding activity, the extracts were first passed over an affinity chromatography column generated with the mutant binding site. The column effluents were collected and the presence of binding activity was confirmed by EMSA. The unbound proteins were then passed over a column of the wild-type binding site to purify the transcription factor. To ensure complete binding, the effluent was passed over the column repeatedly overnight at 4°C using a peristaltic pump. Proteins bound to the wild-type column were then eluted with a 0.1 M step gradient of 0.2-1.0 M KCl. Collected fractions were dialyzed against buffer D and tested for binding activity by EMSA. A typical elution profile of the binding activity is shown in Figure 2. Binding activity eluted between 0.7-1.0 M KCl fractions. No binding activity was detected in the column flow through, suggesting complete removal of binding activity from isolated nuclear proteins.

Fractions showing binding activity were collected, pooled, dialyzed against buffer D and again tested for binding activity using EMSA. Fractions containing binding activity were pooled, and again loaded onto the wild-type oligonucleotide column. Binding activity was eluted with a KCl step gradient as above, and fractions containing binding activity were concentrated on the Centricon C-10 columns (Amicon). As can be seen in Figure 3, concentration of the purified binding activity resulted in an appropriate increase in the per volume activity of the 0.9 M KCl column eluate.

### **Electrophoretic Mobility Shift Assays (EMSA).**

For EMSA, double stranded DNA probes were labeled with Polynucleotide Kinase (NE Biolab, MA) and p32ATP (NEN Dupont, MA). Approximately 10 ug of protein from nuclear extracts was incubated for 30 min in the reaction buffer (20mM HEPES, pH7.9, 20% Glycerol, 100 mM KCl, 0.2 mM EDTA, 1.0 mM PMSF, 1 mM DTT) containing 0.5 nM of 5'-end-labeled double

strand probe, and 1 ug of the nonspecific binding inhibitor, poly dI•dC (Sigma, MO), in a total volume of 15 ul. In competition experiments, a 40-fold excess of the unlabeled competitor oligonucleotide was added to the reaction. The complexes formed were separated on the 6-8% TBE polyacrylamide gels (National Diagnostics) and visualized by autoradiography.

### **Amino acid analysis**

A small amount of concentrated fractions from affinity column were loaded on a 12% SDS-PAGE and gel was subjected to silver staining (Bio-Rad, Hercules, CA). Three separate bands with estimated molecular weights of 37, 38 and 42 kDa were detected and shown to bind specifically to the ets-like response element. The remainder of the purified binding activity in the concentrated fractions was loaded onto a preparative the 10% SDS-polyacrylamide gel and to resolve the three proteins. When the gel was stained with Coomassie blue, two major bands stained with Coomassie blue were identified, indicating that 37 and 38 kDa bands were not resolved on the 10% gel. The two bands corresponding 42 kDa and 38 kDa were cut excised from the gel and washed in acetonitril in preparation for amino acid sequencing. In the original proposal we indicated that we would perform the amino acid analysis in house using the Microsequencing Shared Resource Facility at the Fred Hutchinson Cancer Research Center. Unfortunately turnover of personnel left the facility without a qualified operator during the period when our samples were ready for sequencing. Our previous experience with protein sequencing (5) had included a very successful interaction with the Harvard Microchemistry Laboratory directed by Dr. W. Lane. We therefore decided to use the Harvard facility for amino acid analysis of the larger protein (42kDa). Amino acid sequence analysis was performed using tandem Electro-Spray Mass Spectrometry. The smaller protein 38 kDa protein was subjected to MS/MS analysis at the University of Washington (Seattle) sequencing facility directed by Dr. Aebersold.

Briefly, proteins cut from the preparative gel were digested with Trypsin. The peptide mixture was concentrated on the C18 column to remove salt and hydrophilic molecules before analysis using microcapillary Solid Phase Extraction (SPE)-capillary electrophoresis coupled online with mass spectrometry (LCQ ion trap) (6). Eluting analytes are detected in the MS, and trigger the instrument to switch from MS to MS/MS mode to generate collision-induced dissociated (CID) spectra of automatically selected analyte ions. The CID spectra generated during an experiment are then used to search protein or nucleotide databases using the Sequest software to identify possible sequence matches.



## Results

The use of the sequence specific DNA affinity chromatography proved to highly successful and made it possible to obtain several micrograms of highly purified proteins. During course of protein isolation we found that the binding activity frequently presented as two distinct complexes, often within the same nuclear extract. The relative amounts of the two complexes appeared to vary among individual nuclear preparations, with the larger complex being more prevalent in less dense cell cultures (data not shown). The results of DNA binding competition experiments showed that both binding activities are specific for the *ets*-like response element (Figure 4). To further demonstrate specificity of DNA binding and to estimate the molecular weights of the proteins comprising the different complexes we performed UV cross-linking of proteins to the labeled oligonucleotide as described in Figure 1B of the original proposal. The results indicated that the two complexes detected by EMSA were derived from the binding of two different protein species to the DNA sequence (Figure 5). The molecular weights of the proteins bound to DNA in the larger and smaller complex were estimated by ~42-43 kDa and 37-38 kDa, respectively. The estimate was based on the molecular mass of protein determined by SDS-Page after UV cross-linking minus molecular mass of the single stranded probe, assuming that only one strand was cross-linked per treated complex. As expected, the affinity purified binding activity will be comprised of at least two protein species with molecular weights of 38 and 42 kDa, corresponding the two different cross-linked proteins (Figure 6). On a 12% SDS-PAGE analysis, a third band of 37KDa was also detected by silver staining. The latter protein was not resolved by the 10% preparative gel, and presumably co-migrated with the 38 kDa band.

MS/MS analysis is a powerful new tool for amino acid sequence analysis. The mass of individual peptides fragments generated by CID are compared to a database to identify likely matches in sequence. For the 42 kDa band analyzed by the Harvard Microchemistry lab, a total of 47 polypeptides were positively identified on the basis of their mass (Figure 7 and Table 1). The 21 of these overlapping peptides corresponded to peptides sequence present in a previously described RNA binding protein designated AUF-1(7, 8). Another group of 14 overlapping polypeptides corresponded to those present in a rat transcription factor designated CARG box binding factor or CBF-A (9). An additional 12 overlapping polypeptides corresponded to a region conserved among the heterogeneous nuclear ribonucleoproteins A3, A3(A) and A2/B1. Analysis of the 38 kDa band at the University of Washington facility by tandem mass spectrometry identified 14 overlapping peptides within the AUF-1 and CARG proteins. All peptides identified at University of Washington matched those identified at Harvard Microchemistry Laboratory. Peptides detected by two independent analyses are shown in Table 1. Together the identified peptides detected covered 36% and 37% of the coding sequences of CARG and AUF-1, respectively and are shown in Figure 7.

## Discussion

The finding of peptides corresponding to both the AUF-1 and CArG proteins in both the 38 and the 42 kDa protein bands can be explained in one of two ways. One possibility is that there are actually two proteins present, both proteins are present in multiple co-migrating forms that cannot be resolved by SDS-PAGE. The presence of different forms of the same protein could be the result of differentially spliced and/or post-translationally modified forms of the two proteins that by chance co-migrate. Alternatively, the findings could be the result of cross contamination of the two protein bands, even though they were resolved by 4-5 mm on the preparative SDS-PAGE gel. Further analyses will be required to distinguish between these possibilities. Nonetheless, the sequencing data did suggest two candidate proteins for further study.

Among the two candidate proteins identified, the AUF-1 protein was previously shown to bind specifically to the 3' UTRs (un-translated regions) of the mRNAs encoding c-myc and granulocyte-macrophage colony-stimulating factor genes and stimulate their degradation. Anti-AUF antisera detect two immunologically cross-reactive protein species with molecular weights of 37 and 40 kDa. Both polypeptides are phosphorylated and can be found in a complex(s) with other polypeptides (7). Immunologically related polypeptides were found in both the nucleus and the cytoplasm. AUF-1 protein is about 60% identical to CBF-A transcription factor protein. However, there is no information as to whether or not AUF-1 itself can function as a transcription factor.

The second candidate gene, CArG binding factor or CBF-A protein, is a transcriptional factor that binds to a CArG box sequence and regulates muscle specific activity of the  $\alpha$ -actin and other muscle specific genes. It was also found to regulate the promoters of some early response genes following serum stimulation (10). In cardiac cells, CArG binding protein functions as transcriptional repressor (11). The predicted molecular weight of the CArG binding protein is about 31 kDa (12). However, its migration in SDS-PAGE gels is somewhat anomalous, suggesting a molecular weight of ~41-43 kDa protein (12). A homology search of the protein database identified two additional proteins with extensive homology to the CBF-A protein. These proteins include AIF-C1 and another unnamed nucleic acid binding protein. No information other than amino acids sequence has been published for either of these two proteins. Unfortunately, all of the CBF-A polypeptides identified in our MS/MS analysis are in a highly conserved domain that is 100% identical among these three proteins at the amino acid level (Figure 6). Thus, we cannot distinguish among these proteins at this time.

The CBF-A factor has been shown to have an ability interact with double and single stranded DNA *in vitro*. However, it remains unclear if CBF-A transcription factor is able to interact with single stranded DNA *in vivo* (9). Interestingly, CArG boxes are the sites of DNA bending (13). Torsional stress due to bending can result in local melting of dsDNA, suggesting that CBF-A may bind alternate DNA conformations. CBF-A or closely related proteins are excellent candidates for transcription factor binding to this region. In the original proposal we hypothesized that this region of the promoter was involved in an unusual conformation in mammary cells, and that this conformation affects expression of *Hras1* and other genes with similar promoter elements. Disruption of this conformation by hormones or carcinogens would then permit binding of specific transcription factors and increase gene expression. Although the structure of the CS region in mammary cells remains unclear, it is conceivable that release of the CArG box from the CS conformation would induce DNA bending and allow for binding of CBF-A transcription factor. Consistent with such a model, preliminary data indicate that a CArG box sequence from other genes can compete with specific binding to the *ets*-like sequence present in the CS region of the *Hras1* promoter. It is therefore unlikely that an *ets*-related protein is involved in binding to the CS region of the *Hras1* promoter. Rather a novel mechanisms involving a class of transcription factors that bind to altered DNA conformations seems likely.

### **Future considerations**

Since we identified four candidates (CBF-A, AIF-C1, nucleic acid binding protein and AUF-1) the next step is to verify if proteins are able to interact with the specific element in the *Hras1* promoter. One approach is to determine if the elements to which these proteins bind can compete with specific binding to the *Hras1* promoter element. We are also attempting to obtain antisera to these proteins to determine cross-reactivity with our specific binding activity. We are also attempting to obtain or generate full-length cDNA clones for these proteins that can then be translated *in vitro* and analyzed for binding activity.

Failure to confirm one of the candidates using these approaches would indicate that our specific binding protein differs significantly from these candidate proteins outside the conserved region. In this case the most straightforward approach will be to screen mammary tumor expression library that is available from Invitrogen. This alternative approach was approved in proposal and will let us to identify all possible cDNAs encoding proteins able to bind DNA probes derived from the *Hras* promoter. Since CBF-A was identified by this approach, other members of the same protein family can be identified by similar mean.

## Key Research Accomplishments

- Purification of microgram quantities of specific binding proteins by DNA affinity purification
- Detection and separation of two major protein species with specific binding activity
- Amino acid analysis of two purified proteins by tandem Electro-Spray Mass Spectrometry
- Identification of four candidate genes on the basis of sequence identity with mass spectrometry data – AUF-1 and three highly homologous proteins that may bind alternative DNA conformations

## Reportable Outcomes

- None during the past budget period

## Conclusions.

While *ras* oncogene mutations are not frequently detected in human breast cancers, carcinogen-induced disruption of the CS is unlikely to be an anomaly of the rat. Several lines of evidence suggest that carcinogen-induced CS disruption also play an important role in human breast cancer. Our preliminary data had indicated the presence of an analogous structure within a conserved sequence of the human *Hras* promoter. Other studies have also shown that the *Hras* gene is frequently overexpressed in human breast cancer, suggesting that disruption of the CS in human breast cells could also contribute to the pathogenesis of human breast cancer. There is also no reason to posit that these hormonally regulated DNA structures are restricted to the promoters of *Hras1* genes in mammary cells. Our preliminary studies indicated that the sequences comprising the *Hras1* CS include an *ets*-like responsive element in both the human and rat promoter. The results obtained so far suggest that this promoter element is not an *ets* binding site in mammary cells. Rather, the proteins that bind to this element are likely to be members of a family of proteins that bind to CarG box elements. These proteins may bind to alternative DNA conformations such as bent or single-stranded sequences. Our previous studies indicated that the *Hras1* promoter and other genes can adopt alternative, tissue specific conformations *in vivo*. The presence of these structural variants and proteins that recognize them could therefore play a role in tissue specific regulation of growth related genes such as *Hras*. It is plausible that carcinogen-induced alterations of DNA conformations permit binding of transcription factors such as CBF-A and lead to deregulation of these genes. Further characterization of these binding proteins and their role in regulation of tissue specific gene

expression should provide further insight into their contributions to epigenetic carcinogenesis in both rats and humans.

## References

1. R. S. Cha, W. G. Thilly, H. Zarbl, *Proceedings of the National Academy of Sciences of the United States of America* **91**, 3749-53 (1994).
2. Z. Jin, B. Houle, A. M. Mikheev, R. S. Cha, H. Zarbl, *Cancer Research* **56**, 4927-35 (1996).
3. J. D. Dignam, R. M. Lebovitz, R. G. Roeder, *Nucleic Acids Research* **11**, 1475-89 (1983).
4. J. T. Kadonaga, *Methods in Enzymology* **208**, 10-23 (1991).
5. C. D. Hoemann, H. Zarbl, *Cell Growth & Differentiation* **1**, 581-90 (1990).
6. D. Figeys, Y. Zhang, R. Aebersold, *Electrophoresis* **19**, 2338-47 (1998).
7. W. Zhang, et al., *Molecular & Cellular Biology* **13**, 7652-65 (1993).
8. K. Ehrenman, L. Long, B. J. Wagner, G. Brewer, *Gene* **149**, 315-9 (1994).
9. S. Kamada, T. Miwa, *Gene* **119**, 229-36 (1992).
10. M. Soulez, D. Tuil, A. Kahn, H. Gilgenkrantz, *Biochemical & Biophysical Research Communications* **219**, 418-22 (1996).
11. T. A. Gustafson, L. Kedes, *Molecular & Cellular Biology* **9**, 3269-83 (1989).
12. A. Cvekl, J. B. McDermott, J. Piatigorsky, *Biochimica et Biophysica Acta* **1261**, 290-2 (1995).
13. T. A. Gustafson, A. Taylor, L. Kedes, *Proceedings of the National Academy of Sciences of the United States of America* **86**, 2162-6 (1989).

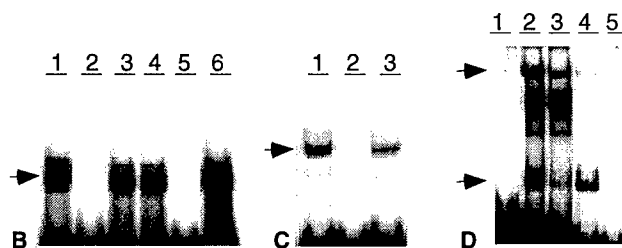
## Appendices

**Table 1.** List of polypeptides identified by two independent MS/MS analyses of the 38 and 42 kDa DNA binding proteins isolated by affinity chromatography.

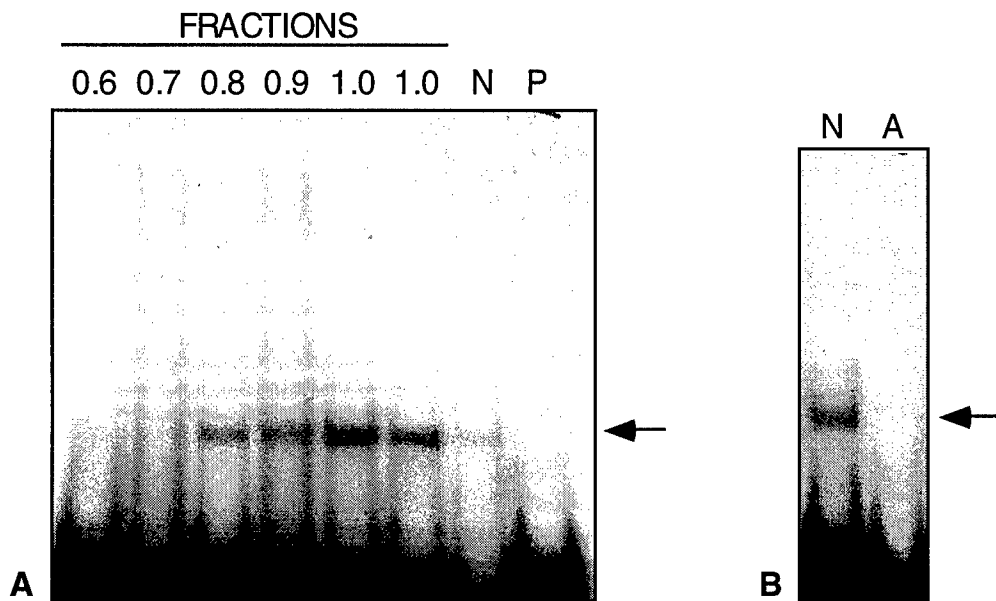
| Polypeptides related to CArG protein | AUF-1 related polypeptides   |
|--------------------------------------|------------------------------|
| MFVGGLSWDTSK                         | IDASKNEEDEGKMFIGGLSWDTTKK    |
| MFVGGLSWDTSKK                        | IDASKNEEDEGKMFIGGLSWDTTK     |
| MFVGGLSWDTSKKDLKDYFTK                | KDLKDYFSK                    |
| DLKDYFTK                             | DLKDYFSK                     |
|                                      | IDASKNEEDEGK                 |
| SRGFGFILFK                           | MFIGGLSWDTTKK                |
| GFGFILFK                             | IGGLSWDTTKK                  |
|                                      | NEEDEGKMFIGGLSWDTTKKDLK      |
| IFVGGLNPEATEEK                       | NEEDEGKMFIGGLSWDTTKKDLKDYFSK |
| IFVGGLNPEATEEKIR                     | DYFSK                        |
| GGLNPEATEEK                          |                              |
| IREYFGQFGEIEAIELPIDPK                | SRGFGFVLFK                   |
| EYFGQFGEIEAIELPIDPK                  | ESESVDKVMQKEHK               |
|                                      | ESESVDKVMQK                  |
| GFVFITFKEEDPVKK                      |                              |
| GFVFITFKEEDPVK                       | IFVGGLSPDTPEEK               |
|                                      | GGLSPDTPEEK                  |
| FHTVSGSK                             | EYFGGFGEVESIELPMDNKTNK       |
|                                      | IREYFGGFGEVESIELPMDNK        |
| EVYQQQQYGSGGR                        | IREYFGGFGEVESIELPMDNKTNK     |
|                                      | KYHNVGLSK                    |

|           |   | Competition |
|-----------|---|-------------|
| RAT HRE   | 5'- TAG GGG <u>TTC</u> CGG <u>CGA</u> ACT CTG T -3' |             |
| MUTANT 1  | _____ T TA A _____                                  | (-)         |
| MUTANT 2  | _____ GCG C _____                                   | (-)         |
| MUTANT 3  | _____ G G _____                                     | (+)         |
| MUTANT 4  | _____ CA _____                                      | (-)         |
| MUTANT 5  | _____ CA _____ G G _____                            | (-)         |
| HUMAN HRE | 5'- TG CTC CCG <u>GAA</u> GCC CC -3'                | (+)         |
| STAT 5/6  | 5'- GTA TTT <u>CCC</u> AGA <u>AAA</u> GGA AC -3'    | (-)         |
| EBS       | 5'- GAT AAC CGG <u>AA</u> G TAA GTA ACG - 3'        | (+)         |

**A**



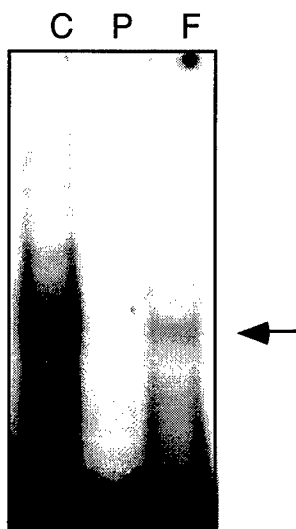
**Figure 1.** Oligonucleotide competition experiments using electrophoretic mobility shift assays demonstrate strong affinity of *ets*-like transcription factors for the *Hras* Responsive Element (HRE) present within the conformation switch sequences from rats and humans *Hras* promoters. Panel **A**) Oligonucleotide sequences of the putative normal and mutant rat HRE (*Hras*-responsive element), the human HRE sequence (see Figure 8 below), and commercially available Stat5/Stat6, and E74 *ets* binding site oligomers. The putative *ets* binding elements are underlined or overlined. Band shift experiments were carried out with labeled rat HRE probe. Non-radioactive competitor oligonucleotides were introduced into the binding reaction in a 40-fold excess. Nuclear extracts (10 ug) from BICR mammary tumor cells were incubated with 0.5 ng of 5' end- labeled rat HRE probe (Panels B, C, and D) or labeled E74 *ets* probe (panel D) and excess competitor DNA for 20 min, resolved on 6% PAGE and analyzed by autoradiogram. The gels used to generate the autoradiograms in panels B, C, and D were not run under identical conditions, and therefore the differences in mobility shifts observed among the panels are of no significance. Arrows indicate the positions of specific bands identified on the basis of competition with excess cold rat HRE probe. Data with mutant oligos 4 and 5 are not shown. Ability of test oligomers to compete with the putative Hras1 HRE for binding of specific transcription factors; Panel **B**) lane 1- no competitor; 2 - human HRE probe; 3 - mutant 1; 4 - mutant 2; 5 - mutant 3; 6 - SP1 element; Panel **C**) Lane 1 - no competitor, 2 - human HRE, 3 - Stat5/Stat6 oligonucleotide. Panel **D**) indicates that the rat HRE and the E74 *ets*-binding site yield similar patterns of gel retardation, as indicated by the formation of two similar complexes (arrows). These results suggested that these two sequences bind the same-*ets* related factors, but with different affinities. Lane 1- labeled E74 probe; lanes 2 - labeled E74 probe; lane 3- labeled E74 probe and excess of rat HRE competitor oligomer; 4 - labeled rat HRE probe; oligomer; line 5 - labeled rat HRE probe and excess competitor E74 oligomer.



**Figure 2. Purification of the ets-like response element binding activity by affinity chromatography.** Panel A. A typical elution profile of the sequence specific binding activity eluted with a 0.1 M step gradient of 0.2-1.0 KCl. Analysis of aliquots from each fraction by EMSA indicated that the bulk of the binding activity was eluted in the 0.8-1.0M KCl fractions (only 0.6-1.0M KCl fractions are shown). Eluted fractions were dialyzed against binding buffer and 10 ul aliquots were incubated with  $^{32}$ P-labeled DNA probe corresponding to the rat promoter (Figure 1) in a presence of poly dI•dC, resolved on 8% native 1x TBE gel, and complexes detected by autoradiography (exposure time was 10 minutes). N- indicates binding activity present in the nuclear protein extract before loading on the column; P- indicates migration position of the labeled DNA probe in the absence of nuclear proteins.

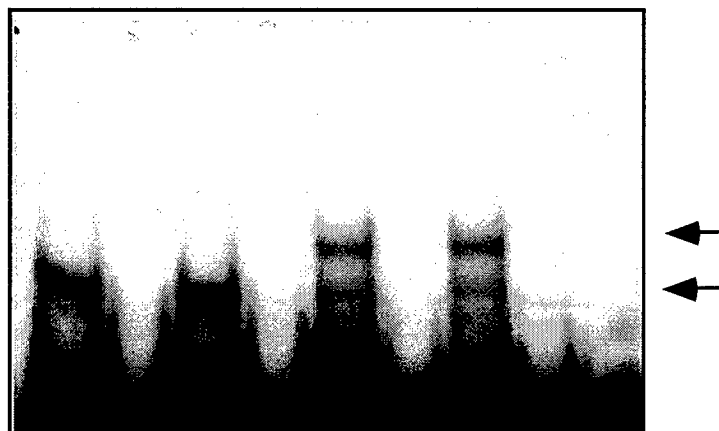
Panel B. Efficient retention of specific binding on the affinity column. Shown is a comparison of specific binding activity in nuclear proteins before (N) and after loading onto the affinity column (F). EMSA analysis was as in Panel A. Exposure time 2.5 hours.



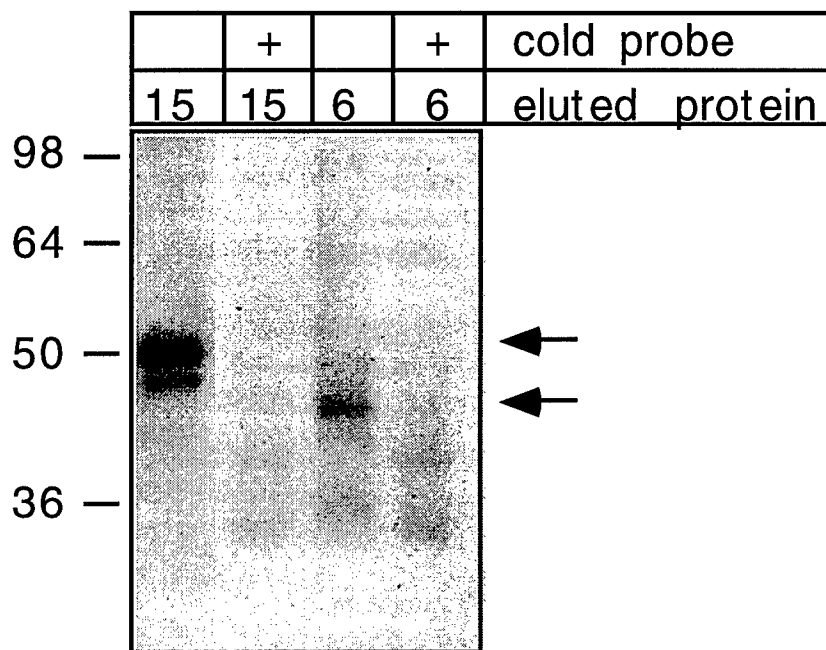


**Figure 3. Retention of specific DNA binding activity after concentration of affinity column purified fractions using Centricon C10 columns.** Affinity column fractions showing significant binding activity were pooled, dialyzed, and again purified by affinity chromatography. The fractions were then eluted with a KCl step gradient, dialyzed and concentrate concentrated from ~1 ml down to 50-80 ul. The concentrated fractions were then reanalyzed for binding activity by EMSA a in Figure 2. Shown in the binding activity of the pooled 0.9 M KCl fraction before and after concentration. **C**- concentrated fraction, **P**-probe without protein added, **F**- fraction before concentration on the Centricon 10 column (Amicon).

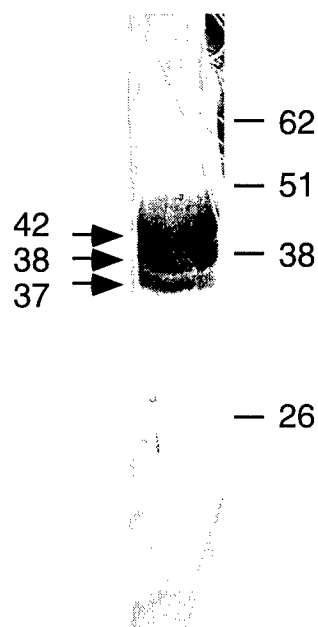
|   |    |   |                |
|---|----|---|----------------|
| 6 | 15 | - | Eluted protein |
| + | +  |   | EBS            |
| + | +  |   | Mutant 1       |
| + | +  |   | Mutant 3       |



**Figure 4. Detection of distinct binding activities in independent nuclear protein preparations.** The variability in the size and relative intensities of specific DNA complexes formed with different nuclear protein preparations are illustrated using Lot 6 and 15 as an example. EMSA analysis of affinity purified binding activity was as in Figure 2. The competition experiment was performed by including a 40-fold excess of cold double strand DNA. The sequences of the wild-type, EBS and mutant DNA oligonucleotides used for the competitions are shown in Figure 1. Arrows indicate the migration positions of the two distinct complexes formed by specific binding of the probes.



**Figure 5. Detection of two distinct species of specific binding proteins by UV cross-linking.** The binding activities detected in affinity purified nuclear proteins from Lot 6 and 15 (shown in Figure 3) were used for UV cross-linking as described in Figure 1B of the original application. Binding reactions were subjected to UV irradiation for 30 min boiled, resolved on the 10% SDS-PAGE and covalent DNA protein complexes detected by autoradiography. Where indicated an excess of unlabeled probe (100 fold) was included in the reaction to demonstrate specificity of cross-linking. The major species detected in Lot 15 predicts a cross-linked protein with a molecular weight of ~42 kDa, while the major species detected in the Lot6 protein preparation predicts a cross-linked protein with a molecular weight of ~38 kDa (see text). Slower migrating species were spurious.



**Figure 6. SDS-PAGE analysis of affinity purified binding proteins.** The affinity purified binding proteins were pooled and concentrated as described in Figure 3. Five percent of the purified sample was loaded onto a 12% SDS-PAGE gel. Silver staining of the resolved proteins was performed with Silver staining Kit (Bio-Rad) according to manufacture protocol. Three bands with molecular weights of ~37, 38 and 42 kDa are detectable by staining. Molecular weight of the proteins was estimated by regression analysis of molecular weight marker distribution (not shown).

|        |  |
|--------|--|
| CARG   | MSDAAEEQPMETTG-----ATENGH-EAAPEGEAPVEPS-AAAAAPAAS-AGSG       |
| AIF-C1 | MSDAAEEQPMETTG-----ATENGH-EAAPEGEAPVEP--SPGSGPATP-AGSV       |
| NABP   | MTWLSQQLRVKCRLLSCWVRLFYFWCCWEAGEVEAVFRCSSELNWSCDYRLKVPNQKKN  |
| AUF1   | MSEEQFGGDGAAAA-----ATAAVGAAAGEQEGAMVAATQGAAAAREADAGPGA       |
|        | *: . * . : . . .   |
|        |  |
| CARG   | GGTTTAPSGNQNGAEGDQINASKNEEDAGKMFVGGLSWDTSKKDLKDYFTKFGEVVDCTI |
| AIF-C1 | GGTTAAPSGNQNGAEGDQINASKNEEDAGKMFVGGLSWDTSKKDLKDYFTKFGEVVDCTI |
| NABP   | GGRGSRPSGNQNG-EGDQINASKNEEDAGKMFVGGLSWDTSKKDLKDYFTKFGEVVDCTI |
| AUF1   | EPRLEAPKGSAES-EGAKIDASKNEEDEGKMFIGGLSWDTTKKDLKDYFSKFGEVVDCTI |
|        | *.*. :. ** :*:***** :*:*****:*****:*****:                    |
|        |  |
| CARG   | KMDPNTGRSRGFGFILFKDSSSVEKVLDQKEHRLDGRVIDPKKAMAMK-KDPVKKIFVGG |
| AIF-C1 | KMDPNTGRSRGFGFILFKDSSSVEKVLDQKEHRLDGRVIDPKKAMAMK-KDPVKKIFVGG |
| NABP   | KMDPNTGRSRGFGFILFKDSSSVEKVLDQKEHRLDGRVIDPKKAMAMK-KDPVKKIFVGG |
| AUF1   | KLDPITGRSRGFGFVLFKESESVDKVMQKEHKLNGKVIDPKRAKAMKTKEPVKKIFVGG  |
|        | *:* *****:***:*:*:*:*:*:*:*:*:*:*:******                     |
|        |  |
| CARG   | LNPEATEEKIREYFGQFGEIEAIELPIDPKLNKRRGFVFITFKEEDPVKKVLEKKFHTVS |
| AIF-C1 | LNPEATEEKIREYFGQFGEIEAIELPIDPKLNKRRGFVFITFKEEDPVKKVLEKKFHTVS |
| NABP   | LNPEATEEKIREYFGQFGEIEAIELPIDPKLNKRRGFVFITFKEEDPVKKVLEKKFHTVS |
| AUF1   | LSPDTPEEKIREYFGGFGEVESIELPMDNKTNKRGFVCFITFKEEEPVKKIMEKKYHNVG |
|        | *.*.:.***** :*:*****:* ***** *****:****:***:*.*              |
|        |  |
| CARG   | GSKCEIKVAQPKEVYQQQQ-YGS-GGRGNRNRGNRSGGGGQ-----               |
| AIF-C1 | GSKCEIKVAQPKEVYQQQQ-YGS-GGRGNRNRGNRSGGGQSQSWNQGYGNYWNQGYGYQ  |
| NABP   | GSKCEIKVAQPKEVYQQQQ-YGS-GGRGNRNRGNRSGGGGQ-----               |
| AUF1   | LSKCEIKVAMSKEQYQQQQQWGSRGGFAGRARGEFRNSSEA-----               |
|        | ***** .** ***** :** ** .*. * *: ...                          |
|        |  |
| CARG   | -----GSTNYGKSQRRGGHQNNYKPY                                   |
| AIF-C1 | QGYGPGYGGYDYSPIGYGYGPGYDYSQGSTNYGKSQRRGGHQNNYKPY             |
| NABP   | -----GSTNYGKSQRRGGHQNNYKPY                                   |
| AUF1   | -----GEGLELPPNSIHCWQLSV---                                   |

Fig 7. Alignment of proteins identified by MS/MS analysis of 37 and 42 kDa proteins isolated by affinity chromatography. Database searches identified four candidate gene on the basis of sequence identity. These include CBP-A, a Car box binding protein (CarG), the AIF-C1 transcription factor, a nucleic acid binding protein (NABP) and the AUF-1 RNA binding protein. Sequences are aligned to show regions of sequence similarity. An \* indicate the position of a conserved amino acid residue. Lines above the alignment show overlapping polypeptides corresponding to CBP-A related proteins. Lines below the alignment show overlapping polypeptides corresponding to AUF-1 protein. Each line corresponds to the total length of the sequence defined by the overlapping polypeptides. Individual polypeptides listed in the table 1. All identified polypeptides were within highly conserved regions.